Supplementary Information:

Supplementary Methods

Summary

Two adult male rhesus monkeys performed a visual attention task, during which they fixated a central spot and released a bar when the behaviourally relevant stimulus underwent a shape change (Extended Data Fig. 2). Behavioural relevance was assigned on a trial-by-trial basis with a centrally presented cue. Two stimuli were presented, one in the lower right visual hemifield, and one in the upper left visual hemifield. Neuronal signals were recorded from the left hemisphere in two monkeys using subdural ECoG grids consisting of 252 electrodes (1 mm diameter), which were spaced 2-3 mm apart[11,15,16]. Data were recorded in 9 sessions in monkey 1 and 14 sessions in monkey 2. The post-cue analysis used the time period from 0.3 s after cue onset until the first shape change in one of the stimuli. Only trials with a correct behavioural report were used. For each trial, this period was cut into non-overlapping 0.5 s data epochs. This resulted in 3874 epochs for monkey 1 and 3492 epochs for monkey 2. For both the pre-stimulus and pre-cue periods, there were 4239 and 4396 epochs of 0.5 s in monkey 1 and 2, respectively. For each site and recording session, the data epochs were normalized by their standard deviation and subsequently pooled across sessions. Data epochs were multitapered using three Slepian tapers and Fourier-transformed[28]. The epoch lengths of 0.5 s resulted in a spectral resolution of 2 Hz, the multitapering in a spectral smoothing of ±3 Hz. Where mentioned explicitly, we used Hann-tapered 1 s epochs for 1 Hz spectral resolution. The Fourier transforms were the basis for calculating the coherence spectra and for calculating the GC influence spectra through non-parametric spectral matrix factorization[19]. The non-parametric estimation of GC influences spectra has certain advantages over parametric approaches, e.g. it does not require the specification of a particular model order.

Experimental paradigm

All procedures for the electrophysiological recordings were approved by the ethics committee of the Radboud University Nijmegen (Nijmegen, The Netherlands). After touching a bar, the acquisition of fixation, and a pre-stimulus baseline interval of 0.8 s, two isoluminant and iso-eccentric stimuli (drifting sinusoidal gratings, diameter: 3 degrees, spatial frequency: ~1 cycles/degree; drift velocity: ~1 degree/s; resulting temporal frequency: ~1 cycle/s; contrast: 100%) were presented on a CRT monitor (120 Hz refresh rate non-interlaced). In each trial, the light grating stripes of one stimulus were slightly tinted yellow, the stripes of the other stimulus were slightly tinted blue, assigned randomly (Extended Data Fig. 2). After a variable amount of time (1-1.5 s in monkey 1, 0.8-1.3 s in monkey 2), the colour of the fixation point changed to blue or yellow, indicating the stimulus with the corresponding colour to be the behaviourally relevant one. A trial was considered correct and the monkey was rewarded when the bar was released within 0.15-0.5 s of the change in the cued stimulus, ignoring equally likely changes of the non-cued stimulus. In monkeys 1 and 2, 94% and 84% of bar releases, respectively, were correct reports of changes in the relevant stimulus. The stimulus change consisted of the stimulus' stripes undergoing a gentle bend, lasting 0.15 s. Either one of the stimuli, irrespective of being cued or not, could change at a random time between stimulus onset and 4.5 s after cue onset. Trials were terminated without reward when the monkey released the bar outside the response window, or when it broke fixation (fixation window: 0.85 degree radius in monkey 1, 1 degree radius in monkey 2). For the analyses presented here, if not specified otherwise, data from all correct trials of both attention conditions were pooled.

Neurophysiological recordings

The ECoG grids were implanted under aseptic conditions with isoflurane/fentanyl anaesthesia. Intra-operative photographs were acquired for later coregistration. Signals were amplified, low-pass filtered at 8 kHz and digitized at roughly 32 kHz with a Neuralynx Digital Lynx acquisition system. Local Field Potentials were obtained by low-pass filtering at 250 Hz and downsampling to 1 kHz.

Data analysis general

Data analysis used the FieldTrip toolbox (Supplementary ref.[31]). If not stated otherwise, all analyses used bipolar derivations, i.e. sample-by-sample differences between immediately neighbouring electrodes. Bipolar derivations were used in the analysis only when both electrodes had been assigned to one and the same area, excluding pairs of electrodes that crossed area boundaries. As a control for the arbitrary absolute phases obtained from bipolar derivations, we also used a current-source density (CSD). See figure S1 for further details on bipolar derivations and CSD. Power line artefacts at 50, 100 and 150 Hz were estimated and subtracted from the data using a Discrete Fourier Transform (DFT).

We defined individual beta and gamma bands in each monkey by using a peak detection algorithm that fitted parabolas to the coherence spectrum averaged across all site pairs of all ROIs (Extended Data Fig. 3) and taking the full width at half maximum. For the post-cue period, this resulted in the following bands: in monkey 1, the gamma band was 67-83 Hz (peak frequency was 74Hz), and the beta band was 12-24 Hz (peak frequency was 18 Hz). In monkey 2 the gamma band was 54-74 Hz (peak frequency was 64 Hz), and the beta band was 7-21 Hz (peak frequency was 14 Hz). The theta band was defined on individual peaks of the coherence spectrum averaged across all site pairs of all ROIs and taking half of the maximum as width. This resulted in the following theta band for both monkeys: 2-6 Hz (peak frequency was 4 Hz). In both monkeys, the gamma, beta and theta band peaks were the only reliably detected peaks in these average spectra. The same method was applied for the definition of individual beta and gamma bands in the other periods of the task, the pre-stimulus period from fixation to stimulus onset, and the pre-cue period from stimulus onset until cue presentation. This gave nearly identical results, except for the period preceding stimulus onset, where no gamma peak could be detected in the average spectra and correspondingly, the gamma band was not included in Fig. 3**c** and Extended Data Fig. 7.

Analysis of conditional Granger-causal influences

For the analysis of conditional GC influences, the bipolar derived signals were first low-pass filtered with a cutoff at 90 Hz and downsampled to 300 Hz. Subsequently, the dimensionality was reduced by performing a principal component analysis (PCA) on the time courses of all signals from a given area and keeping only the principal components (PCs) that explained most variance until at least 90% of the variance of that area was explained. Across all areas, in monkey 1, 97 sites were reduced to 73 PCs, reducing the number of pairs entered into the non-parametric spectral matrix factorization from 4656 to 2628; in monkey 2, 108 sites were reduced to 80 PCs, reducing the number of pairs from 5778 to 3160.

We computed block-wise, conditional GC influences between each pair of areas, treating the PCs representing all other areas as the block to be conditioned on[21]. Consider the PCs belonging to area 1, to area 2, and to the remaining areas. To compute the conditional GC influence that area 1 exerts onto area 2, conditioned on the rest, we performed multivariate non-parametric spectral matrix factorization on two cross-spectral density matrices: 1) on the cross-spectral density matrix containing all PCs, 2) on the cross-spectral density matrix containing PCs from area 2 and the remaining areas to be conditioned on. The resulting transfer functions and noise covariance matrices from the two factorizations are used to derive the GC influence from area 1 onto area 2, conditioned on the rest, which quantifies, per frequency, the unique variance in area 1 that contributes to predictions about area 2, above and beyond the variance present in the other areas. This procedure was repeated for all possible pairs of areas, in both directions.

Analysis excluding microsaccade effects for the gamma band

Horizontal and vertical eye position was monitored at 230 Hz. Microsaccades (MSs) were detected using a velocity threshold of 5 standard deviations. We selected all pairs of MSs that were separated by at least 0.8 s. Of those 0.8 s, we discarded 0.3 s post-MS and used the remaining 0.5 s for the analysis. At 0.3 s after a microsaccade, the LFP gamma phase is no longer locked to the microsaccade, and it is generally not phase-locked to an upcoming microsaccade[22]. Thus, these 0.5 s epochs were used to analyse GC influences in the absence of any effects from gamma locking to microsaccades.

Region of Interest (ROI) definition

For both monkeys implanted with ECoG grids, individual structural MRIs were acquired and the brains were segmented. ECoG electrode positions were co-registered with the segmented brains based on high-resolution intra-operative photographs, using the sulci for alignment (Extended Data Fig. 1). In order to assign an electrode to a cortical area, we co-registered the individual segmented brains to the F99 template brain (CARET v5.62). On the F99 brain, several different monkey brain atlases are defined (Supplementary ref.[32-34]): “Felleman and Van Essen.1991 (FVE91)”, “Paxinos et al.2000 (PHT00)”, “Markov et al.2010 (CC10)” and “Markov et al.2012 (CC12)”. These atlases were projected onto the individual segmented brains (Fig. S1). Thereby, for each atlas, each ECoG electrode was assigned to a cortical area. Across atlases, an electrode was assigned to the area to which it was assigned in the majority of atlases. If there was a tie, we considered a fifth atlas, not available in CARET, namely the atlas by Saleem and Logothetis (Supplementary ref.[35]). With regard to 7A, one of the atlases (“Paxinos et al.2000”) distinguishes between the more lateral 7A/PG and the more medial 7A/OPT. The majority of 7A studies concerned with visual function have dealt with 7A/OPT (Supplementary ref.[36-39]), and also the 7A retrograde tracer injection used here, targeted the medial part of 7A (Supplementary ref.[33]), i.e. was most consistent with an injection in 7A/OPT. Therefore, we restricted our definition of 7A to 7A/OPT and defined the lateral boundary of 7A according to “Paxinos 2000”. After final electrode assignment, we selected the electrodes assigned to the areas V1, V2, V4, TEO, 8L, 8M, DP, 7A (Fig. 1 and Extended

Data Fig. 1). Bipolar derivations were used in the analysis only when both electrodes had been assigned to one and the same area, excluding pairs of electrodes that crossed area boundaries. The resulting sites were distributed as follows: Monkey 1: V1: 31 sites, V2: 9 sites, V4: 19 sites, DP: 10 sites, TEO: 5 sites, 8/FEF: 15 sites (8M: 7 sites; 8L: 8 sites), 7A: 8 sites; Monkey 2: V1: 50 sites, V2: 14 sites, V4: 18 sites, DP: 8 sites, TEO: 3 sites, 8/FEF: 5 sites (8M: 2 sites; 8L: 3 sites), 7A: 10 sites.

Retrograde tracer database

Description of the anatomical dataset acquisition and analysis has been reported in Markov et al., 2012 (Supplementary ref.[33]). The values that we used correspond to multiple injections each into V1, V2, V4 and single injections into areas DP, TEO, 8/FEF (8L and 8M) and 7A. SLN values were obtained as described in Fig. 3a[4]. For the correlation with DAI values, only SLN values based on at least 10 labelled neurons were included. Updates, atlases and additional information concerning the anatomical dataset that was used for this work is available at www.core-nets.org.

Statistical testing

We first tested for each area pair, whether the average GC influence between all inter-areal site pairs was significant, i.e. whether it reliably exceeded the bias level. We estimated the bias by randomly pairing epochs before GC influence calculation. For each of 500 randomizations, the mean over the GC influences in the two directions was placed into a randomization distribution and the 95th percentiles of the resulting distributions were used to determine the bias level. Every inter-areal GC influence reported in Extended Data Figures 5, 6 and 7 exceeded the bias level.

For a given directed influence and a given frequency band, we used the bootstrap method to estimate the standard error and the 95% confidence intervals in order to determine whether the GC influences in the bottom-up and top-down directions were significantly asymmetric (Supplementary ref.[40]). Confidence intervals and the resulting statistics are reported in figures 1, 4 and extended data figures 3c-d, 5, 6, 7, 8 and 10.

In the analysis shown in figure 2b, we used the Pearson correlation coefficient as the first-level statistic, and based the statistical inference on a non-parametric randomization approach with multiple comparison correction at the cluster level as described in supplementary reference[41]. To this end, we repeated the following procedure 1000 times: 1) We randomly paired the DAI and SLN values. 2) We calculated the DAI-SLN correlation coefficient as the first-level statistic. 3) We determined the frequencies, whose first-level statistic was significant at $p<0.05$. 4) We searched for clusters of contiguous frequencies with significant correlation and calculated the mean correlation coefficient per cluster as the second-level statistic. 5) We placed the largest second-level statistic into the randomization distribution. For the non-randomized DAI-SLN correlation values, we also performed steps 2) through 4). Clusters observed in the non-randomized data were considered significant if their second-level statistic was smaller than the 2.5th percentile or greater than the 97.5th percentile of the randomization distribution, corresponding to a two-sided test with a threshold at $p<0.05$, corrected for the multiple comparisons across frequencies.

All other reported correlation coefficients are based on the Spearman rank correlation. This pertains to the following figures and/or the corresponding text: Fig. 2c,d,e, Fig. 3a,b,c, Extended Data Fig. 9a,b,c,d.

Supplementary Discussion

The absolute values of inter-areal coherence and GC influence

The absolute values of inter-areal coherence and GC influence are small. The theta-, beta- and gamma-band peaks of inter-areal coherence averaged over all inter-areal site pairs fell into the range of 0.04 to 0.1, the respective GC influences into the range of 0.0015 to 0.005 (Extended Data Fig. 3). The coherence values show that on average about 0.16 to 1% of the variance in the local theta-, beta, or gamma-band activity is explained by the respective activities in any one of the sites in the other areas. The GC influence values show that most of this explained variance can in turn be attributed to Granger-causal influences.

These amounts of explained variance might well be highly relevant, which becomes clear if one compares the physiology to the anatomy. When a cortical target site is injected with retrograde tracer, the labelling reveals the neurons providing direct synaptic inputs to the injected site. Of those input neurons, 70-90% stem from the injected area (Supplementary ref.[33] – Figure 2). The remaining 10-30% of inputs distribute across the other source areas, such that a given typical source area contributes only a small fraction of the synaptic inputs to a given target site. This fraction ranges over roughly 6 orders of magnitude from less than 0.00005% up to approximately 19%. Across the area pairs studied here, this fraction ranged from 0.00023% (for V1 input to 7A) to 19% (for V2 input to V1), with a median of 0.05%. If inside a source area, one picked a random site, this site would contribute again only a fraction of all the inputs deriving from this source area.

This strongly suggests an anatomical basis for the small fractions of local theta-, beta-, and gamma-band activity that are explained by the respective activities in any one of the sites in the other areas: The small fractions of explained variance are most likely due to the small fractions of synaptic input deriving from the respective sites. Yet, even if the fraction of synaptic input to a target site from a particular source site is small, it is a prime structural candidate mechanism for inter-areal cortico-cortical communication. By the same token, even if the fraction of variance explained by coherence is small, coherence is a prime candidate process for inter-areal communication.

Supplementary References

31 Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J. M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience* **2011**, 156869, doi:10.1155/2011/156869 (2011).

32 Markov, N. T. *et al.* Weight consistency specifies regularities of macaque cortical networks. *Cereb Cortex* **21**, 1254-1272, doi:10.1093/cercor/bhq201 (2011).

33 Markov, N. T. *et al.* A Weighted and Directed Interareal Connectivity Matrix for Macaque Cerebral Cortex. *Cereb Cortex*, doi:10.1093/cercor/bhs270 (2012).

34 Van Essen, D. C. Windows on the brain: the emerging role of atlases and databases in neuroscience. *Current opinion in neurobiology* **12**, 574-579 (2002).

35 Saleem, K. S. & Logothetis, N. K. *A combined MRI and histology atlas of the rhesus monkey brain in stereotaxic coordinates*. (Academic Press, 2007).

36 Constantinidis, C. & Steinmetz, M. A. Neuronal responses in area 7a to multiple stimulus displays: II. responses are suppressed at the cued location. *Cereb Cortex* **11**, 592-597 (2001).

37 Constantinidis, C. & Steinmetz, M. A. Neuronal responses in area 7a to multiple-stimulus displays: I. neurons encode the location of the salient stimulus. *Cereb Cortex* **11**, 581-591 (2001).

38 Rawley, J. B. & Constantinidis, C. Effects of task and coordinate frame of attention in area 7a of the primate posterior parietal cortex. *Journal of vision* **10**, 12 11-16, doi:10.1167/10.1.12 (2010).

39 Raffi, M. & Siegel, R. M. Functional architecture of spatial attention in the parietal cortex of the behaving monkey. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **25**, 5171-5186, doi:10.1523/JNEUROSCI.5201-04.2005 (2005).

40 Efron, B, & Tibshirani, R,J.An introduction to the bootstrap. *Chapman & Hall/CRC Monographs on Statistics & Applied Probability*, pp1-465, isbn:0-412-04231-2 (1994).

41 Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *Journal of neuroscience methods* **164**, 177-190, doi:10.1016/j.jneumeth.2007.03.024 (2007).